

CLAIMS

What is claimed is:

1. A system that facilitates online printing for a web user comprising:
an analysis component that analyzes data corresponding to a plurality of web pages, the plurality of web pages comprising at least a current web page, to facilitate identifying at least one section of the current web page; and
a parsing component that dissects the current web page to remove at least one identifiable section from printable view, the at least one section being non-essential to defining content of the current web page.
2. The system of claim 1, further comprising a preview component that visualizes a preview version of the printable page for the web user.
3. The system of claim 2, the preview component allows the web user to modify the printable page according to user preferences.
4. The system of claim 2, the preview component allows the web user to verify whether the parsing component dissected the current web page according to user preferences.
5. The system of claim 1, the plurality of web pages further comprises at least one web page in addition to the current web page to facilitate deciphering architecture of the current web page.
6. The system of claim 1, the analysis component analyzes at least one of text density, key word density, key word frequency, table structure, location of dense text, location of one or more table, presence of one or more images, location of one or more images, dimensions of one or more images, and dimensions of one or more tables.

7. The system of claim 1, the parsing component identifies one or more sections of the current page based at least in part upon a portion of the data from the current web page that has been analyzed by the analysis component.

8. The system of claim 7, the parsing component labels one or more identified sections on the current web page with one or more visual references to assist the web user in determining whether to remove one or more identified sections from the printable view.

9. The system of claim 7, the parsing component labels one or more identified sections on the current web page with one or more textual references to assist the web user in determining whether to remove one or more identified sections from the printable view.

10. The system of claim 1, further comprising a crawler component that provides data for analysis by the analysis component.

11. The system of claim 10, further comprising a cache component operatively connected to the crawler component, the cache component comprising one or more web pages stored for comparison to one or more updated pages that correspond thereto to facilitate determining location of content.

12. The system of claim 1, further comprising a feed monitoring component that monitors data from one or more feeds to facilitate deciphering content on at least the current web page for analysis by the analysis component.

13. The system of claim 1, non-essential data comprising advertisements and navigation information.

14. The system of claim 1 further comprising a user selection component that allows the web user to select which identifiable sections to keep or remove from the printable view.

15. The system of claim 1, the printable view is a preview version of a print copy of the current web page.

16. The system of claim 1, the printable view corresponding to any one of following as chosen by the web user: one or more sections selected by the parsing component to be retained for printing, the current web page as it appears, and one or more sections of the current web page fit to a number of pages according to user preferences.

17. A system that facilitates providing a preview version of a printable page comprising:

an attribute analysis component that analyzes data from one or more web pages, the one or more web pages comprising at least a current web page and at least one other web page; and

an artificial intelligence (AI) component that performs a utility-based analysis in connection with analyzed data to automatically generate the preview version based at least in part upon user preferences.

18. The system of claim 17, the AI component comprising a classifier.

19. The system of claim 17, the AI component inferring at least one of the following:

when to employ user preferences;

when to override user preferences;

when to compare data from the current web page and the at least one other

web page; and

when to remove one or more sections from the preview version.

20. The system of claim 17, the AI component comprising at least one of: a support vector machine (SVM), a naïve Bayes model, a Bayesian network, a decision tree, a Hidden Markov Model (HMM), neural network, data fusion engine.

21. A method that determines web page architecture to facilitate identifying content on a web page comprising:

providing a plurality of web pages; and

analyzing the plurality of web pages to facilitate identifying content and determining location of the content on at least a subset of the web pages to facilitate generating a user-defined printable page.

22. The method of claim 21, the plurality of web pages comprising at least a current web page and at least one other web page.

23. The method of claim 22, further comprising printing the current web page as it appears on screen.

24. The method of claim 21, analyzing the plurality of web pages for at least one of the following to facilitate identifying content and determining location of the content: text density, key word density, key word frequency, table structure, location of dense text, location of one or more table, presence of one or more images, location of one or more images, dimensions of one or more images, and dimensions of one or more tables.

25. The method of claim 21, further comprising determining layout structure of at least a subset of the web pages based at least in part on content identified thereon.

26. A method that facilitates printing one or more web pages comprising: analyzing a plurality of web pages for their respective web structures to facilitate identifying content, the plurality of web pages comprising at least one current web page;

dissecting the at least one current web page based at least in part upon identifiable content to into one or more sections that are subject to removal before printing; and

generating a preview version of a printable page to a web user, the preview version having one or more sections removed therefrom.

27. The method of claim 26, further comprising modifying the preview version based at least in part upon user preferences.

28. The method of claim 26, dissecting the current web based at least in part upon user preferences.

29. The method of claim 26, further comprising collecting one or more current web pages and organizing the one or more current web pages, thereby generating a new printable document comprising at least a subset of the current web pages based at least in part upon user preferences.

30. The method of claim 26, further comprising printing the printable page to at least one of the following: to file, to printer, to email system, and to facsimile device.

31. The method of claim 26, identifiable content comprising advertisements, images, navigation, body of text, dynamic text.

32. The method of claim 26, adding a thumbnail reference to the preview version comprising an image of the current web page in its original state.

33. The method of claim 26, adding a page summary to the preview version comprising a URL and a print date.
34. A print preview page of a web page comprising:
 - at least a subset of content parsed from a web page; and
 - at least one of the following: a thumbnail reference of the web page and a summary of the web page.
35. The print preview page of claim 34, the content is user-defined content.
36. The print preview page of claim 34, the summary comprising a URL of the web page and a date on which the web page is accessed.
37. A data packet adapted to be transmitted between two or more computer processes facilitating providing suggestions to an online user, the data packet comprising:
 - information associated with analyzing a plurality of web pages for their respective web structures to facilitate identifying content, the plurality of web pages comprising at least one current web page; dissecting the at least one current web page based at least in part upon identifiable content into one or more sections that are subject to removal before printing; and presenting a preview version of a printable page to a web user, the preview version having one or more sections removed therefrom.
38. A computer-readable medium having stored thereon the following computer executable components:
 - an analysis component that analyzes data corresponding to a plurality of web pages, the plurality of web pages comprising at least a current web page, to facilitate identifying at least one section of the current web page; and
 - a parsing component that dissects the current web page to remove at least one identifiable section from printable view, the at least one section being non-essential to defining content of the current web page.

39. A system that facilitates printing from a web page comprising:
 - means for providing a plurality of web pages; and
 - means for analyzing the plurality of web pages to facilitate identifying content and determining location of the content on at least a subset of the web pages to facilitate generating a user-defined printable page.

40. The system of claim 39, further comprising means for determining layout structure of at least a subset of the web pages based at least in part on content identified thereon.